Artificial Intelligence-based Cognitive Cross-layer Decision Engine for Next-Generation Space Mission

Anu Jagannath
Advanced Applied Technology
ANDRO Computational Solutions, LLC
Rome, NY, United States
ajagannath@androcs.com

Jithin Jagannath
Advanced Applied Technology
ANDRO Computational Solutions, LLC
Rome, NY, United States
jjagannath@androcs.com

Andrew Drozd
Advanced Applied Technology
ANDRO Computational Solutions, LLC
Rome, NY, United States
adrozd@androcs.com

***Abstract*—In this position paper, the authors argue the need for a novel framework that provides flexibility, autonomy and optimizes the use of scarce resources to ensure reliable communication during next-generation space missions. To this end, the authors present the shortcomings of existing space architectures and the challenges in realizing adaptive autonomous space-networking. In this regard, the authors aim to jointly exploit the immense capabilities of deep reinforcement learning (DRL) and cross-layer optimization by proposing an artificial intelligence-based cognitive cross-layer decision engine to bolster next-generation space missions. The presented software-defined cognitive cross-layer decision engine is designed for the resource-constrained Internet-of-Space-Things. The framework is designed to be flexible to accommodate varying (with time and location) requirements of multiple space missions such as reliability, throughput, delay, energy-efficiency among others. In this work, the authors present the formulation of the cross-layer optimization for multiple mission objectives that forms the basis of the presented framework. The cross-layer optimization problem is then modeled as a Markov Decision Process to be solved using deep reinforcement learning (DRL). Subsequently, the authors elucidate the DRL model and concisely explain the deep neural network architecture to perform the DRL. This position paper concludes by providing the different phases of the evaluation plan for the proposed cognitive framework.**

***Index Terms*—Artificial intelligence, deep reinforcement learning, cross-layer optimization, cognitive decision engine, internet-of-space-things, software defined radios.**

I. Introduction

As NASA enters a new era of space exploration, where communication links shift from point-to-point communications to networked topologies involving relay satellites, spacecraft swarm, nanosatellites (CubeSats), multiple robotic vehicles communicating with each other and with ground terminals, certain communication bottlenecks emerge. These include spectrum congestion, low rate links, energy-unaware communication at the physical layer, unreliable resource-negligent networking among others. Conventional legacy systems that adopt fixed rate communications, layer-specific approaches, static routing, etc. cannot address these communication challenges. There have been significant layer-specific efforts to enhance space communication such as variable rate communication using adaptive modulation and coding [1], [2] at physical layer, variants of traditional routing protocols for space communications at network layer. However, these approaches are specific to the layers in the protocol stack. The benefit gained by sharing information between layers – *cross-layer approach* – has been under study for terrestrial and space wireless communication systems recently but are not yet in practice for space communication [3]–[14].

In recent years several small (micro-, nano-/CubeSats, and pico-) satellites, TETwalkers, and Tracking and Data Relay Satellites (TDRS) have been deployed, which communicate with national and international ground stations for research missions, planetary surface exploration, among others. Several other Internet-of-Space-Things (IoSTs) such as Marsbees, Mars Helicopter, among others are envisioned to be deployed in space especially for Mars surface exploration. With the advent of these communicating entities, communication can extend from long point-to-point links to hop-by-hop communication links by means of wireless networking as in Fig. 1. The different types of wireless links that exist in space communication are intra-vehicle, inter-vehicle, planetary surface-to-surface, planetary surface-to-spacecraft and space-to-earth. All these wireless links except the intra-vehicle can range from several kilometers to several hundreds or thousands of kilometers. A point-to-point link for such long-range communication implies higher transmission power to attain reliable communication. A hop-by-hop distributed networking approach implies significant onboard power savings and reliable com-

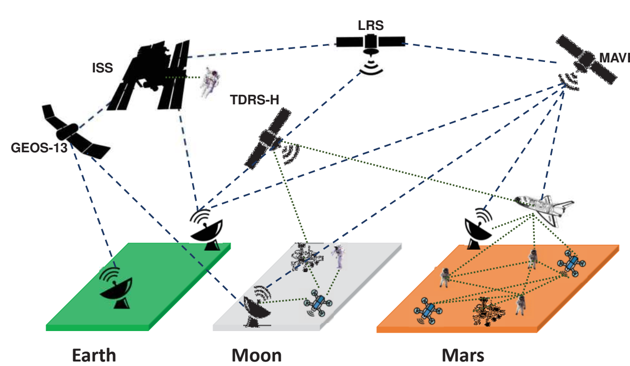

Fig. 1. Space Network

munication considering the resource-constrained architecture of these IoSTs. Therefore, in this work, we aim to discuss the challenges of existing space communication architectures and argue in favor of the immense performance benefits that can be attained by designing an artificial intelligence (AI)-based cognitive cross-layer decision engine. Accordingly, we describe the need for such a cognitive engine, provide the initial design of the framework and present implementation plan to drive this position paper forward.

The paper is organized as follows. Section II discusses the existing space communication architecture, challenges and opportunities. Section III will detail the framework of the proposed AI-based cross-layer decision engine. Section IV discusses the implementation plan. Finally, we summarize our conclusions in Section V.

II. Related Works and Motivation

A. Existing Architecture.

Traditional satellite communications have been point-to-point and *bent-pipe* in nature. Such a bent-pipe approach is typical of Low Earth Orbit (LEO) satellites which are designed to communicate directly with the ground terminal or via a TDRS [15]. The inter-satellite distance in LEO networks can range from ten to several thousands of kilometers resulting in propagation delays in the order of milliseconds that is significantly higher than terrestrial networks [16]. The authors of [17] overcome this long distance link problem by satellite-terrestrial segment networking which integrates satellites with inter-satellite link (ISL) capability and ground terminals to form nodes of a large ad hoc network. GAMANET brings in the flexibility of software-defined radios (SDRs) to realize this space-to-earth ad hoc network. Another non-agile feature of traditional satellite networks is the *slow configuration* whereby the ground terminals configure the satellites when it passes directly over the terminal. This results in an *inflexible approach* which involves static routing, link allocation and scheduling that lack the intended agility to support the varying fading, traffic effects and quality of service (QoS) requirements. Consequently, current satellite communication cannot sustain dynamic fine-grained QoS guarantee.

With the advent of small satellite systems in the nano, pico and micro range of satellites, the notion of space-based networking is becoming increasingly tangible. Few of the small satellite missions are ANTS (NASA's Autonomous NanoTechnology Swarm), GRACE (joint venture by NASA and Deutsche Forschungsanstalt fr Luftund Raumfahrt in Germany), EDSN (NASA's Edison Demonstration of SmallSat Network), PROBA-3 (Small satellite demonstration mission by European Space Agency), QB-50 etc. Small satellite networks allow for higher reconfigurability, unified mission objective, scalability at a reduced cost with respect to larger satellites. The formation flying aspects introduce swarm, cluster, trailing and constellation formations. The network topology of satellites at a particular time instant is referred to as the topology slice. The topology slice varies with time as satellites come in and go out of range in a satellite network. The varying topology, link quality, frequency availability, mission-specific QoS requirements pose significant challenges to the inter-satellite communication in small satellite networks as well as large formations. Satellite resources such as storage, power and processing are limited and expensive. These constraints are further restricted in the case of small satellite systems.

Traditional satellite systems possess layered architecture typical of the Open Systems Interconnection (OSI) reference model. Such a modular protocol stack approach independently optimizes parameters in the respective layers leading to redundancies and inefficiencies. The QoS requirements are serviced at the upper layers although they are affected by the lower level protocols. Lack of scalability and adaptability with the fluctuating network dynamics can result in significant performance inefficiencies. Inefficient allocation of radio resources, for instance, larger number of time slots to a satellite that is currently experiencing fading effects lead to higher error rates. This inefficiency stems from lack of information exchange between link and physical layer resulting in wasteful resource allocation. Several strategies [18]–[20] have been proposed to adaptively vary the modulation and coding scheme for DVB-S2 links with varying channel conditions such that higher order modulation schemes are chosen at good channel conditions and lower order modulations are chosen when channel quality is poor. Since the physical layer does not consider the end-to-end data rate requirement for the current application being serviced, this could lead to excessive delays for high data rate applications. Such inefficiencies arising from lack of information exchange between the layers of protocol stack can be better addressed with a unified cross-layer approach.

B. Cross-layer approach

A cross-layer approach is imperative to address the inefficiencies arising from the layered architecture and the unique challenges in realizing space networking. Ensuring optimal radio resource allocation, medium access and routing strategies while satisfying QoS requirements involve interaction with the disparate layers of the protocol stack. The cross-layer design for terrestrial ad hoc networks have been studied extensively [3]–[9] but its application to space networks is still at its infancy with very few notable works in this direction [10]–[14].

The authors of [3] propose a cross-layer communication framework to ensure QoS for heterogeneous applications in wireless multimedia sensor networks. The work in [5] illustrates the use of a unified optimization framework as applied to opportunistic scheduling in single-hop cellular networks and joint congestion-control and scheduling problem in multihop wireless networks. The authors of [6] analyzed the effect of imperfect scheduling on cross-layer congestion control for wireless networks. Cross-layer optimized protocols have been playing a vital role in maximizing the performance of various ad hoc networks and to mitigate the unique challenges of modern communication networks. Due to the critical nature of data flowing through tactical ad hoc networks, a deadline based cross-layer routing protocol was proposed in [7] to maximize the effective throughput of the network. The effectiveness of such an approach was demonstrated by evaluation on SDR-based cross-layer testbed [7], [8]. Similarly, an energy-aware routing protocol was proposed in [9] to bolster emergency ad hoc public safety network and was shown to achieve twice the network lifetime of traditional shortest path approaches. Several of these optimization objectives can be applicable to space mission but usually involves much more adaptable parameters due to the heterogeneous nature of the IoSTs. Hence, demanding the need to employ deep learning solution to ensure the feasibility of next-generation cross-layer architectures for space exploration.

There has been some effort to apply cross-layer techniques for satellite communication. The authors of [13] proposed a cross-layer bandwidth allocation scheme which involves physical and link layers. The scheme adopts a master-slave model for satellite network whereby the master allocates bandwidth to the slaves based on their bandwidth requests and local channel conditions. The cross-layer information from physical layer and link layer is used for dynamic uplink packet scheduling based on channel conditions for unicast DVB-S2 services in [11]. A MODCOD adaptation at the physical layer based on transport layer goodput performance is proposed in [14]. The authors assume a GEO bent-pipe model forming a relay link between two TCP client/server-to-earth stations where the transmission mode adaptation is performed by the transmitting earth station based on transport layer ACKs from the server at the receiving end. A dynamic resource allocation scheme for DVB-RCS satellite networks based on cross-layer interaction between the transport and link layer is proposed in [10]. The authors consider a GEO stationary satellite receiving control signaling from a network control center (NCC) in the ground terminal and serving multiple user terminals based on the control parameters as dictated by the NCC. Although these works apply cross-layer technique for satellite communication, *their scalability and adaptability as applied to a satellite constellation or spacecraft swarm may be intractable as the network grows.* Primarily because the decision making ability are not distributed to the network entities (satellites/spacecrafts) and relies on a control entity located either in the ground or chosen among the satellite formation (swarm/cluster/constellation) which dictates the actions to be adopted by each satellite in the formation. Extending such centralized approaches to large satellite formations are impractical and uneconomical considering the resource-constrained expensive space missions. A potential cross-layer solution for next-generation space-missions must, therefore, be autonomous, scalable (decentralized), and reconfigurable based on changing mission objectives. Maintaining such distributed cross-layer framework would demand handling large number of operational parameters across multiple layers of the protocol stack. To handle this parametric explosion, we exploit the powerful artificial intelligence (AI) tool called *Deep Learning*. Therefore, in this work, we present a novel communications framework that employs cross-layer technology and deep learning to provide an autonomous and adaptive solution to enable next-generation space missions.

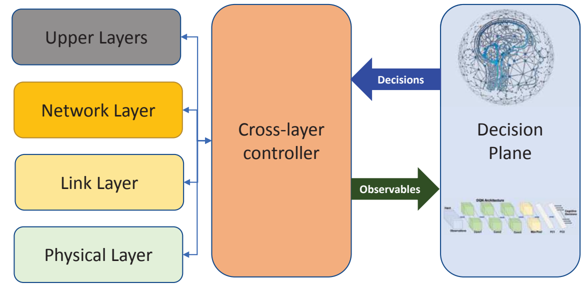

Fig. 2. Cognitive Cross-layer Decision Framework

III. Cognitive Cross-layer Decision Framework

In this section, we present the cognitive cross-layer decision framework shown in Fig. 2 designed to enable efficient communication and networking ability for space exploration. Let's assume the space ad hoc network model with multiple IoSTs and/or terrestrial terminals forming the nodes of the network. The end-to-end QoS metrics are derived from the upper layers (such as transport layer) of the protocol stack. The routing decisions are derived from the network layer and the transmission resources are assigned and configured at the physical layer. In the proposed framework, the cross-layer controller will gather multiple parameters such as transmission power, bit rate, frequency band, residual battery energy, packet destination, queue backlog, QoS requirements among others from multiple layers of the protocol stack. The gathered parameters are then made available to the decision engine to make optimal decisions.

In this work, we consider two different scenarios to present a flexible framework that can accommodate multiple mission objectives. For instance, let us suppose the mission objectives are (i) QoS-aware energy efficient networking and (ii) resource-aware reliable networking. With respect to each mission goal, the framework derives multiple layer-specific objectives that will eventually be jointly attained by the cross-layer optimization.

Physical Layer Formulation. From the perspective of physical layer, the objective is to perform optimum radio resource allocation in response to the varying link dynamics and spectrum occupancy. Let us consider N number of nodes in the network with F number of available frequency bands to communicate with each other. The set of available nodes can be denoted as $\mathcal{N}$. The maximum onboard transmission power can be denoted by $\mathcal{P}$ and let P_f^n represent the transmission power of node n on the frequency band f. Each user performs spectrum sensing on the allocated communication band and returns the interference plus noise metric that is piggy-backed to control signaling to immediate neighbors such that they can use the obtained link metrics in optimization. The interference plus noise metric of node n on the F frequency bands is denoted by the vector $\mathbf{i}^n = [i_1^n, i_2^n, \cdots, i_F^n]$. The energy efficiency (η_f^{nm}) of a transmission from the node n to m can be defined as the number of bits successfully transmitted per Joule of transmission energy in frequency band f at a transmission bit rate of r^n. Now, the score metric can be derived as,

$$score_1 = \eta_f^{nm} = \frac{\left(1 - e_f^{nm}\right) r^n}{P_f^n}, \quad (1)$$

where e_f^{nm} (function of i_f^m) is the bit error rate (BER) of link between n and m at frequency band f. Now, the first optimization objective can be derived as,

$$\begin{aligned} &\textbf{Objective}_1\textbf{: Maximize } score_1 \\ &\textbf{Find: } \left\{f, r^n, P_f^n\right\} \\ &\textbf{Subject to: } r^n \geq r^{QoS}, P_f^n < \mathcal{P}, \qquad \forall n \in \mathcal{N}, \forall f \quad (2) \\ &e_f^{nm} \leq e^{QoS} \qquad \forall n, m \in \mathcal{N}, \forall f \quad (3) \end{aligned}$$

Here, the constraints (2) and (3) guarantee the acceptable QoS level of the application being serviced by ensuring; the transmission bit rate is not less than the minimum, r^{QoS} as requested by the service, and BER is below the acceptable threshold e^{QoS}. The constraint $P_f^n < \mathcal{P}$ ensures the transmission power stays within the bound. The transmission bit rate is affected by the modulation and coding scheme (MODCOD) chosen. Hence, choosing a certain bit rate selects the appropriate MODCOD scheme.

Network Layer Formulation. The two different architectures considered for instance are the small satellite constellation and swarm. In the first instance, the space ad hoc network involves small satellite constellation communicating with few ground stations. In the second case, we consider spacebot swarms such as Marsbees or Mars Helicopters communicating with each other and a rover that acts as the data aggregation center. In both cases, the entities of the ad hoc network will be referred to as nodes with the destinations (ground station/rover) acting as the data sink (S). We represent this space network as a directed connectivity graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$ where $[\mathcal{L}(n, m)] \in \mathcal{E}\ \forall m \in \mathcal{N}_n$ is the set of wireless links in the network. Here, $\mathcal{N}_n$ is the set of potential next-hops of n. The distance between any nodes n and m is given as d^{nm} and their queue backlogs are represented as q^n, and q^m. The residual and initial battery energies of node m can be denoted as b_r^m and b_i^m.

To achieve the mission goal *A: QoS aware energy efficient networking*, the second objective must be formulated to maximize network lifetime in support of extended duration space missions. Network lifetime can be defined as the time until which the first node in the network exhausts its battery power. The networking strategy adopted in this work follows a distributed routing whereby each node (n) with a packet to transmit to a destination S will find the optimal next-hop by considering the forward progress, differential queue backlog $\Delta q^{nm} = q^n - q^m$, and its residual battery energy. Thus, score metric to achieve mission goal *A* is

$$score_2^A = \left(\frac{\max\left[\Delta q^{nm}, \upsilon\right]}{q^n}\right)\left(\frac{d^{nS} - d^{mS}}{d^{nS}}\right)\left(\frac{b_r^m}{b_i^m}\right), \quad (4)$$

where υ is a very small positive constant value to avoid negative value in the first term of equation (4). The $score_2^A$ is computed only when the queue backlog q^n is non-zero, i.e., when the node has outbound packets in the buffer. Similarly, to achieve mission goal *B: resource-aware reliable networking*, the second objective must be formulated to maximize reliability and throughput in a resource-aware manner. To this end, we introduce a distributed route reliability metric (ι^{mS}) of the potential next-hop m to S as in [21] that commensurates the probability of successfully delivering a packet from n to S on the first attempt and the queue backlog of each session. The route reliability metric of a node signifies the reliability of the route to a destination through a given next hop. The route reliability metric varies with topology and link reliability and hence is time variant. The score to achieve goal *B* can be derived as,

$$score_2^B = \iota^{mS}\left(\frac{d^{nS} - d^{mS}}{d^{nS}}\right). \quad (5)$$

The second optimization objective can be obtained as,

$$\begin{aligned} &\textbf{Objective}_2\textbf{: Maximize } score_2^{\{A,B\}} \\ &\textbf{Find: } m \\ &\textbf{Subject to: } \text{Mission-specific constraints.} \quad (6) \end{aligned}$$

Here, $score_2^{\{A,B\}}$ implies the mission specific $score_2$ metric which could be $score_2^A$ or $score_2^B$. Next, we look at how the physical layer and network layer scores will be reformulated to obtain a unified cross-layer score.

Cross-layer optimization The proposed framework employs a cross-layer controller to interact with the multiple layers of the protocol stack. The optimization objectives derived above will be jointly optimized to achieve the dictated mission goal. Therefore, the weighted joint score can be expressed as,

$$score_x = w_1 score_1 + w_2 score_2^{\{A,B\}} \quad (7)$$

where the weights are user-defined as per the mission requirements such that $w_1 + w_2 = 1, \forall\ w_1, w_2 \in (0, 1)$. Now, the joint optimization objective can be expressed as,

$$\begin{aligned} &\textbf{Objective}_x\textbf{: Maximize } score_x \\ &\textbf{subject to: } \text{Combined QoS and resource constraints} \end{aligned}$$

The parametric space will grow as the node's potential next-hop neighbors and their tunable parameters as per the mission objective across the protocol stack layers increase. This is the motivation behind using Deep Reinforcement Learning (DRL) as it combines the advantages of reinforcement learning (RL) with the powerful aspect of Deep Neural Network (DNN).

A. Deep Learning

Deep Learning is a subject of growing interest to researchers from various domains. The applicability of machine learning in wireless communication has been studied in [22], [23]. However, deep learning in conjunction with the cross-layer approach for space communication has not been substantially explored. In this work, we show a succinct presentation of how these two technologies can unite to benefit the space networking aspect for next-generation space missions. Further, *autonomy* is a much-needed aspect for next-generation space missions owing to the intelligent and decentralized operation resulting from it. In this regard, the cross-layer optimization problem discussed above can be formulated as a Markov Decision Process (MDP) and solved using DRL. MDP is modeled as a set of states and actions such that the system is rewarded for taking a certain action. The objective of a MDP is to find the optimal policy that will maximize the cumulative discounted reward

$$\Gamma = \sum_{i=0}^{\infty} \gamma^{i} \tau_{t+i} \tag{8}$$

where $\gamma \in [0,1]$ is the discount factor and τ_{t+i} is the instantaneous reward. The state represents an abstraction of the environment the agent makes decisions on. From the cross-layer perspective, state represents multiple parameters derived from the various layers of the protocol stack. Consequently, the state of the node at time instant t can be represented as

$$s_t = (\mathbf{e}_t, \mathbf{r}_t, \mathbf{f}_t, \mathbf{P}_t, \mathbf{i}_t, \mathbf{q}_t, \mathbf{d}_t, \mathbf{b}_{rt}, \mathbf{b}_{it}, \boldsymbol{\eta}_t, \boldsymbol{\iota}_t, \mathbf{N}_t) \tag{9}$$

where,

$$\mathbf{e}_t = \left[e_t^{nm}\right]^{F\times|\mathcal{N}_n|}, \mathbf{i}_t = \left[i_f^{m}\right]^{F\times|\mathcal{N}_n|}, \boldsymbol{\eta}_t = \left[\eta_f^{nm}\right]^{F\times|\mathcal{N}_n|} \tag{10}$$

States in equation (10) represent the BER, interference measure and energy efficiency of the potential next-hops in the available frequency bands respectively.

$$\mathbf{r}_t = [0, 0, \cdots, r^n, 0, \cdots, 0]^{1\times F}, \tag{11}$$

$$\mathbf{P}_t = \left[0, 0, \cdots, P_f^n, 0, \cdots, 0\right]^{1\times F}. \tag{12}$$

Here, $\mathbf{r}_t$ and $\mathbf{P}_t$ denote transmission bit rate and power of the node and has zeros everywhere except the occupied channel index.

$$\mathbf{q}_t = [q^n, q^m]^{1\times|\mathcal{N}_n|+1}, \mathbf{d}_t = \left[d^{nS}, d^{mS}\right]^{1\times|\mathcal{N}_n|+1}, \tag{13}$$

$$\mathbf{b}_{rt} = [b_r^n, b_r^m]^{1\times|\mathcal{N}_n|+1}, \mathbf{b}_{it} = [b_i^n, b_i^m]^{1\times|\mathcal{N}_n|+1}, \tag{14}$$

while states in equation (13) signify the queue backlog and proximity to S of potential next-hops. The states in equation (14) denote the residual and initial battery energies of the potential next-hops respectively.

$$\mathbf{f}_t = [f]^{1\times F}, \mathbf{N}_t = [m]^{1\times|\mathcal{N}_n|}, \boldsymbol{\iota}_t = \left[\iota^{mS}\right]^{1\times|\mathcal{N}_n|} \tag{15}$$

Finally, states in equation (15) represent the available frequency bands, potential next-hops, and distributed route reliability metrics of potential next-hops respectively. The action at time step t can be denoted as $\mathbf{a}_t = \left(r^n, f^n, P_f^n, m\right)$ that decides the optimum transmission parameters and next-hop. We adopt the model-free approach to learn the optimal policy. The MDP policy maps state to action as $a_t = \pi(s_t)$. Let us denote the total number of possible transmission bit rates as per the MODCOD options as R and available discrete transmission power levels as T. Now, the total number of possible actions are $A = R\times T\times F\times|\mathcal{N}_n|$. The agent attempts to maximize the reward Γ. Since the goal of the cross-layer optimization is to maximize the cumulative score $score_x$, we will formulate the reward as

$$\tau_t = \begin{cases} score_x & \text{if constraints are satisfied,} \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Now, the Q-function of MDP can be computed as,

$$\mathcal{Q}(s_t, a_t) \leftarrow \mathcal{Q}(s_t, a_t) + \lambda\left[\tau_{t+1} + \gamma \max_{a_{t+1}} \mathcal{Q}(s_{t+1}, a_{t+1}) - \mathcal{Q}(s_t, a_t)\right] \tag{17}$$

where $\lambda \in (0,1)$ is the learning rate. The optimal policy is

$$\pi^*(s_t) = \arg\max_{a} \mathcal{Q}(s, a) \tag{18}$$

The DNN being used to solve the DRL cross-layer problem posed above is referred to as the Deep Q-Network (DQN). The proposed DQN will adopt the convolutional neural network (CNN) to extract the useful features from the large state-space. The input state-space will be reformulated as an image tensor $\Lambda(s_t)$ to fully exploit the CNN. The DQN will comprise three convolutional layers, a max-pooling layer, and two fully-connected layers with sigmoid activation functions as in Fig. 3.

Another design challenge we foresee, is the instability of traditional RL when a non-linear approximator such as a neural network is used to represent the Q-function. To overcome this, we adopt the experience replay and target network to improve Q-network stability. During the training process, the DQN $(\mathcal{Q}(\Lambda(s), a, \chi))$ with parameters χ maps the input observed states to actions. During each mapping, the DQN generates a history tuple consisting of state $\Lambda(s_t)$, action a_t, next state s_{t+1}, and reward τ_{t+1}. The history tuple is stored into a replay memory of size M. The target network $\widetilde{\mathcal{Q}}$ with parameters $\widetilde{\chi}$ is copied from the DQN every K steps. At each time step, a minibatch of size M_{mini} is sampled from the replay memory. For each experience tuple sampled from the replay memory, the loss is calculated with respect to target $\widetilde{\mathcal{Q}}$ value ξ_t as,

$$\mathfrak{L}(\chi) = \mathrm{E}\left[\xi_t - \mathcal{Q}(\Lambda(s_t), a_t, \chi_t)\right] \tag{19}$$

$$\xi_t = \begin{cases} \tau_{t+1} & \text{if } a_t \text{ meets constraints,} \\ \tau_{t+1} + \gamma \max_{a_{t+1}} \mathcal{Q}(\Lambda(s_{t+1}), a_{t+1}, \widetilde{\chi}) & otherwise. \end{cases} \tag{20}$$

Batch-normalization will be adopted to accelerate the training process.

IV. Evaluation Plan

The proposed framework will employ the agility of SDRs to realize the envisioned outcomes. First, the framework in itself will be extensively simulated/emulated in MATLAB/GNURadio to establish proof-of-concept. Subsequently, the framework will be implemented in C++/Python for actual testing at the ANDRO' *Marconi-Rosenblatt Applied AI Research Laboratory* fitted with various SDRs such as the Universal Serial Radio Peripheral N-series, X-Series, and the

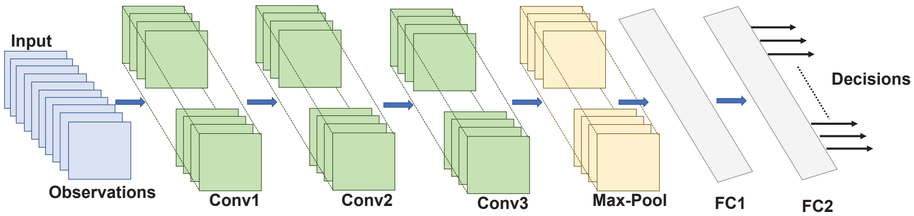

Fig. 3. Deep Q-Network Architecture

Fig. 4. Heterogenous SDRs for Testbed

Zedboard SDR (Fig. 4) among others [8]. The goal is to leverage the heterogeneity of available SDRs to prove its seamless integration on disparate SDR platforms intended for near-future space exploration.

The final phase of testing will potentially employ a software-defined satellite constellation to operate as an ad hoc space network. To demonstrate the resource allocation and adaptive space networking, we will run a test script to automatically attenuate, virtually modify other available resources on the nodes. The testing script will involve multiple test cases to demonstrate the cognitive ability of the framework that runs on each node in a distributed fashion.

V. Conclusion

This position paper puts forth the vision to develop an AI-based cognitive cross-layer decision engine to enhance the existing space communication architecture. Accordingly, the need to unify a cross-layer approach with deep learning has been motivated by discussing the shortcomings of the state-of-the-art and requirements of next-generation space exploration. To this end, a multi-objective cross-layer optimization problem was framed and modeled as an MDP. Subsequently, it has been shown how DRL is chosen as the cognitive entity to overcome the large state-space associated with the cross-layer optimization. Keeping in mind the upcoming resource-constrained IoSTs, the proposed software-defined cognitive framework is aimed to provide the required autonomy and reconfigurability to allow seamless future upgrades and accommodate space missions with multiple/varying mission goals.

References

[1] D. A. Nugroho, M. R. Khaefi, and D. Kim, “Dynamic rate adaptation for industrial WLAN,” in *Proc. of International Conference on ICT Convergence (ICTC)*, pp. 575–580, Oct 2013.

[2] M. Song, C. Xin, Y. Zhao, and X. Cheng, “Dynamic spectrum access: from cognitive radio to network radio,” *IEEE Wireless Communications*, vol. 19, pp. 23–29, Feb 2012.

[3] T. Melodia and I. F. Akyildiz, “Cross-layer QoS-aware Communication for Ultra Wide Band Wireless Multimedia Sensor Networks,” *IEEE J.Sel. A. Commun.*, vol. 28, pp. 653–663, June 2010.

[4] I. Al-Anbagi, M. Erol-Kantarci, and H. T. Mouftah, “A Survey on Cross-Layer Quality-of-Service Approaches in WSNs for Delay and Reliability-Aware Applications,” *IEEE Communications Surveys Tutorials*, vol. 18, pp. 525–552, Firstquarter 2016.

[5] X. Lin, N. B. Shroff, and R. Srikant, “A tutorial on cross-layer optimization in wireless networks,” *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 1452–1463, Aug 2006.

[6] X. Lin and N. B. Shroff, “The impact of imperfect scheduling on cross-layer congestion control in wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 14, pp. 302–315, April 2006.

[7] J. Jagannath, S. Furman, T. Melodia, and A. Drozd, “Design and Experimental Evaluation of a Cross-Layer Deadline-Based Joint Routing and Spectrum Allocation Algorithm,” *IEEE Transactions on Mobile Computing*, pp. 1–1, 2018.

[8] J. Jagannath, H. Saarinen, T. Woods, J. O’Brien, S. Furman, A. Drozd, and T. Melodia, “COmBAT: Cross-layer Based Testbed with Analysis Tool Implemented Using Software Defined Radios,” in *Proc. of IEEE Conf. on Milt. Commun. (MILCOM)*, (Baltimore, MD, USA), Nov 2016.

[9] J. Jagannath, S. Furman, A. Jagannath, L. Ling, A. Burger, and A. Drozd, “HELPER: Heterogeneous Efficient Low Power Radio for Enabling Ad Hoc Emergency Public Safety Networks,” *Ad Hoc Networks (Elsevier)*, vol. 89C, pp. 218–235, 2019.

[10] P. Chini, G. Giambene, D. Bartolini, M. Luglio, and C. Roseti, “Dynamic resource allocation based on a TCP-MAC cross-layer approach for DVB-RCS satellite networks,” *International Journal of Satellite Communications and Networking*, vol. 24, pp. 367–385, Sept 2006.

[11] F. Vieira, M. A. V. Castro, and G. S. Granados, “A tunable-fairness cross-layer scheduler for DVB-S2,” *International Journal of Satellite Communications and Networking*, vol. 24, pp. 437–450, Sept 2006.

[12] N. Celandroni, F. Davoli, E. Ferro, and A. Gotta, “Networking with multi-service GEO satellites: cross-layer approaches for bandwidth allocation,” *International Journal of Satellite Communications and Networking*, vol. 24, pp. 387–403, Sept 2006.

[13] N. Celandroni, F. Davoli, E. Ferro, and A. Gotta, “Adaptive cross-layer bandwidth allocation in a rain-faded satellite environment,” *International Journal of Communication Systems*, vol. 19, pp. 509–530, June 2006.

[14] S. Kota, G. Giambene, and N. L. Candio, “Cross-layer approach for an air interface of GEO satellite communication networks,” *Intl. Journal of Satellite Comm. and Networking*, vol. 25, pp. 481–499, Sept 2007.

[15] W. D. Ivancic, “Architecture and system engineering development study of space-based satellite networks for nasa missions,” in *NASA Technical Reports Server*, 2003.

[16] R. Radhakrishnan, W. W. Edmonson, F. Afghah, R. M. Rodriguez-Osorio, F. Pinto, and S. C. Burleigh, “Survey of Inter-Satellite Communication for Small Satellite Systems: Physical Layer to Network Layer View,” *IEEE Communications Surveys Tutorials*, vol. 18, pp. 2442–2473, Fourthquarter 2016.

[17] P. Rodrigues, A. Oliveira, R. Mendes, S. Cunha, R. Pinho, C. Salotto, and R. De Carvalho, “GAMANET: Disrupting communications and networking in space,” vol. 6, pp. 4237–4245, Jan 2013.

[18] J. A. Downey, M. A. Evans, and N. S. Tollis, “DVB-S2 Experiment over NASA’s Space Network,” in *NASA Technical Reports Server*, 2017.

[19] J. Downey, D. Mortensen, M. Evans, J. Briones, and N. Tollis, “Adaptive coding and modulation experiment with nasa’s space communication and navigation testbed,” in *NASA Technical Reports Server*, 2016.

[20] J. A. Downey, D. J. Mortensen, M. A. Evans, and N. S. Tollis, “Variable Coding and Modulation Experiment Using NASA’s Space Communication and Navigation Testbed,” in *NASA Technical Reports Server*, 2016.

[21] J. Jagannath and T. Melodia, “VL-ROUTE: A Cross-Layer Routing Protocol for Visible Light Ad Hoc Network,” in *Proc. of IEEE Symp. on a World of Wireless, Mobile, and Multimedia Networks (WoWMoM)*, (Washington D.C., USA), June 2019.

[22] J. Jagannath, N. Polosky, A. Jagannath, F. Restuccia, and T. Melodia, “Machine learning for wireless communicationsin the internet of things: A comprehensive survey,” arXiv, 2019 [under review].

[23] C. Zhang, P. Patras, and H. Haddadi, “Deep learning in mobile and wireless networking: A survey,” *IEEE Communications Surveys Tutorials*, pp. 1–1, 2019.